Navy Personnel Research and Development Center

San Diego, CA 92152-6800

TN 90-1 October 1989

(2)

Refinement of the Naval Reserve Officer Training Corps (NROTC) Scholarship Selection Composite

NPRDC TN 90-1 October 1989

Refinement of the Naval Reserve Officer Training Corps (NROTC) Scholarship Selection Composite

Cynthia K. Owens-Kurtz
Walter C. Borman
Kathleen A. Gialluca
Norman M. Abrahams
Personnel Decisions Research Institute
Minneapolis, MN 55414

Joyce D. Mattson
Navy Personnel Research and Development Center
San Diego, CA 92152-6800

Reviewed by
James A. Riedel

Approved by
Patricia J. Thomas

Released by
Richard C. Sorenson
Director, Personnel Systems Department (Acting)

Navy Personnel Research and Development Center
San Diego, California 92152-6800

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
October 1989

3. REPORT TYPE AND DATE COVERED
Technical Note--Apr 87-Jun 88

4. TITLE AND SUBTITLE
Refinement of the Naval Reserve Officer Training Corps (NROTC) Scholarship Selection Composite

5. FUNDING NUMBERS
0602233N, RM33M20

6. AUTHOR(S)
C. K. Owens-Kurtz, W. C. Borman, K. A. Gialluca, N. M. Abrahams, J. D. Mattson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Navy Personnel Research and Development Center
San Diego, California 92152-6800

8. PERFORMING ORGANIZATION REPORT NUMBER
TN 90-1

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Technology (Code 222)
800 North Quincy Street
Washington, DC 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The main purpose of the present study was to evaluate the Quality Index (QI), a weighted predictor composite intended to predict performance in the Navy ROTC program. The QI utilizes high school rank, Scholastic Aptitude Test (SAT) scores a Navy officer interview rating, and Strong Campbell Interest Inventory (SCII)/Background Questionnaire (BQ) keys to predict NROTC student performance. In previous research (Mattson, Neumann, & Abrahams, 1987), a linear composite of these five predictors was found to correlate to a useful degree with school performance as measured by grade point average (GPA), naval science grades (NSG), and ratings of aptitude for performing as a Navy officer (APT). The present study examined the original QI, developed on the 1979 and 1980 entering classes, and other composites of the five predictors on samples of NROTC students from the 1982 and 1985 entering classes. Results suggested that the QI's validity holds up quite well across the samples, with cross validities corrected for restriction-in-range reaching .30-.35 for some of the performance criteria. A second major research activity involved attempting to enhance the effectiveness of individual predictor components of the QI. Accordingly, first, we prepared a new conversion table for transforming ACT test scores to their SAT equivalents, resulting in less overall measurement error than when the existing table is employed. Second, the predictive validity of average and highest SAT or SAT-equivalent composites were compared. The highest composite was found to be slightly more predictive of NROTC performance. Third, we suggested alternative sources of the high school rank predictor in cases where these data are missing. Fourth, suggestions were made to explore new procedures for the officer interview. Finally new empirical keys for the SCII and BQ were developed and then validated on a separate sample. Results showed that these keys improved prediction of NROTC performance when combined with SAT, high school rank, and officer interview predictors.

14. SUBJECT TERMS
Personnel selection. U.S. Navy, NROTC, college students, Navy officer, officer performance

15. NUMBER OF PAGES
39

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

FOREWORD

In 1987 the Chief of Naval Education and Training (CNET) began using a selector composite developed by this Center to predict an applicant's likely academic and military performance in the Navy Reserve Officer Training Corps (NROTC) scholarship program.

This report describes the development and evaluation of an improved selector composite that has also been adopted for use by CNET (N-11). Although NROTC students are already of superior caliber, the use of this composite is expected to result in further improvement in quality.

The work was performed under contract (N66001-D-0085) by the Personnel Decisions Research Institute. The contracting officer's technical representative was Dr. Joyce Mattson.

The work was conducted under the sponsorship of the Office of Naval Technology (Code 222) within the exploratory development Program Element 0602233N, Project RM33M20.

RICHARD C. SORENSON
Director, Personnel Systems Department (Acting)

SUMMARY

Problem

The Navy Reserve Officer Training Corps (NROTC) is a major source of Navy officers. Consequently, selecting the most qualified applicants to receive scholarships is of great concern. Previous research developed and validated the Quality Index (QI), an equation used to linearly weight relevant predictors of performance. The validity of the QI for predicting performance of students in subsequent NROTC entering classes has not been fully assessed, however.

Objectives

The objectives of the present research were to: (1) determine the validity of the QI in subsequent samples, and (2) improve the usefulness of individual components of the QI.

Approach

The QI was applied to students from the 1982 and 1985 entering classes who had complete first-year performance criteria (N = 1452 and N = 1329, respectively), and to students from 1982 who had complete second-year criteria (N = 1202). To develop new ACT-SAT conversion tables, equipercentile equating was applied to data from 1985 (N = 3092) applicants who had taken both ACT and SAT entrance exams. The new tables were then evaluated on 1985 and 1982 (N = 3122) applicants who had taken both exams. The zero-order validities of average versus highest SAT or SAT-equivalent composites were computed on 1983 students who had complete second-year performance criteria and complete SAT or ACT data (N = 773). New empirical keys for the Strong Campbell Interest Inventory (SCII) and the Biographical Questionnaire (BO) were developed in the 1982 samples and cross-validated on students from 1983 who had complete second-year performance criteria (N = 845).

Results and Discussion

In general, the validity of the QI for predicting success in NROTC remains stable across entering classes and criterion years. The new ACT-SAT conversion tables are more accurate than the tables currently in use. The highest SAT or SAT-equivalent composite is slightly more predictive of NROTC performance than is the average composite. Previous research and other concerns (e.g., equity) may warrant the use of the average composite, however. Self-reported high school rank should be investigated as an alternative when these data are missing. A structured, construct-oriented officer interview is suggested to replace the current officer interview. The newly developed SCII and BQ keys improve prediction of NROTC performance when combined with SAT, high school rank, and officer interview predictors. These incremental validities are slightly higher than the incremental validities obtained with the existing SCII and BQ keys, which were designed to predict career retention rather than NROTC performance.

Recommendations

It is suggested that:

1. A close approximation of the QI should be used to aid in the selection of NROTC scholarship recipients.

2. The new ACT-SAT conversion tables should replace the tables currently in use.

3. Although highest SAT/ACT score is slightly more valid than average score, previous research and other concerns (e.g., equity) warrant the use of the average score for selection of Navy ROTC applicants.

4. The value of using self-reported high school rank to replace missing data on this variable should be investigated.

5. A new structured officer interview should be developed.

6. Although the new SCII and BQ keys provide slightly greater incremental validity than the existing keys, no recommendation is warranted until further research determines their relationship to retention.

7. A new BQ should be developed and validated.

CONTENTS

LIST OF TABLES

INTRODUCTION

Problem and Background

The Navy Reserve Officer Training Corps (NROTC) is a major source of Navy officers. Approximately 40,000 individuals apply for NROTC scholarships each year. Scholarships are offered to roughly 3,500 of these applicants. Selection of scholarship recipients is based in part on scores on several predictors: verbal and mathematics scores on the Scholastic Aptitude Test (SAT), high school record rating (essentially, the class rank), interviewers' ratings, Strong Campbell Interest Inventory (SCII) career retention scale score, and Background Questionnaire (BQ) career retention score. These scores are combined in a general linear equation to optimally predict success in NROTC. This equation, called the Quality Index (QI), was developed on the 1979-1980 entering NROTC classes (Mattson, Neumann & Abrahams, 1987). In addition to the QI, the selection board considers leadership potential, high school extra-curricular activities, special circumstances that might affect high school performance, and several other factors in choosing scholarship recipients.

As characteristics of the applicant pool change, it is possible that the QI will require revision. The present research was undertaken to determine the validity of the QI on two subsequent entering classes, 1982 and 1985. In addition, other methods of enhancing the selection procedures for the NROTC program were investigated.

Objectives

This study has several objectives: (1) verify the predictive validity of the QI in samples other than that on which it was developed; (2) evaluate and revise the conversion table used to convert ACT scores to equivalent SAT scores; (3) for applicants who take multiple SAT/ACT examinations, investigate the feasibility of using the mean rather than highest score as input to the QI; (4) discuss alternatives for dealing with missing high school rank; (5) discuss possibilities for improving the officer interview; and (6) investigate new SCII and BQ predictor keys.

APPROACH

Overview

The validation strategy used approximates a triple cross-validation. QI weights developed on 1979-1980 entering classes were applied to 1982 and 1985 entering classes. In addition, regression weights were developed to predict 1982 first-year (1985F), 1982 second-year criteria (1982S), and 1985 first-year (1985F) criteria. Each set of weights was then applied to the non-development samples. Results from such cross-validation provide a stable estimate of the predictive validities of the regression equations. Figure 1 describes this validation approach.

New ACT-SAT conversion tables for verbal, math, and verbal + math composite scores were created through equipercentile equating on data from entry year 1985. These tables were evaluated on 1985 and 1982 applicants against the conversion table currently in use.

Cross-validation Sample	Development Sample 1979-80	1982F	1982S	1985F
1979-80	V			
1982F	C	V	C	C
1982S	C	C	V	C
1985F	C	C	C	V

Note. V = Validation; C = Cross-validation.

Figure 1. Overview of cross-validation design for prediction equations.

The feasibility of using mean versus highest SAT or ACT score was investigated by obtaining validities of highest SAT or SAT-equivalent and mean SAT or SAT-equivalent against college performance criteria using the 1983 second-year (1983S) sample.

New SCII and BQ scoring keys were investigated by examining the relationships of items with criteria in the 1982F and 1982S samples. Separately for the SCII and the BQ, those items related to selected performance-in-college criteria (such as grade point average) were chose to form scoring keys. The incremental validities of the new keys, when combined with SAT, high school rank, and officer interview predictors, were then assessed on the 1983S sample.

Samples

Tables 1 and 2 describe the race and gender distributions, respectively, of the 1979-1980, 1982F, 1982S, 1983S, and 1985F samples. Individuals in these samples had complete predictor and criterion data, were Navy option, and received a 4-year national competition scholarship. BOOST program participants were excluded. Nonrigidities represented approximately 95 percent of each sample. Males represented between 97 percent and 98 percent of each sample. In terms of race and gender, the samples were quite similar.

Measures

1. SAT or SAT-equivalent (SAT): ACT English and mathematics scores were converted to equivalent SAT scores using the conversion table developed by Neumann (1978). For each individual, the SAT or SAT-equivalent verbal and mathematics scores were added to create a composite SAT score. In the case of multiple examinations, the individual's score was the highest composite received. Composite scores were used because equal weights were desired for the verbal and mathematics subtests.

2. High School Record Rating (HSR): HSR is a value approximately equal to the individual's percentile rank in high school class.

Table 1

Race Distributions

Race	1979-80 Freq	1979-80 %	1982F Freq	1982F %	1982S Freq	1982S %	1983S Freq	1983S %	1985F Freq	1985F %
Nonminority	2042	94.9	1389	95.7	1151	95.8	800	94.7	1257	94.6
Minority	110	5.1	63	4.3	51	4.2	45	5.3	72	5.4
Total	2152	100.0	1452	100.0	1202	100.0	845	100.0	1329	100.0

Table 2

Gender Distributions

Gender	1979-80 Freq	1979-80 %	1982F Freq	1982F %	1982S Freq	1982S %	1983S Freq	1983S %	1985F Freq	1985F %
Male	2083	96.8	1427	98.3	1177	97.9	826	97.8	1301	97.9
Female	69	3.2	25	1.7	25	2.1	19	2.2	28	2.1
Total	2152	100.0	1452	100.0	1202	100.0	845	100.0	1329	100.0

3. Interviewer's Ratings (INTVIEW): Each applicant is interviewed by a Navy officer who rates the applicant's potential as a career Navy officer on a five-point scale, with one low and five high.

4. SCII Career Retention Score (SCRET): Neumann and Abrahams (1978) developed this scoring key to predict retention of Navy officers one year beyond required duty.

5. BQ Career Retention Score (BQRET): This index was designed to predict retention of Navy officers 2 years beyond required service.

Dependent Variables

Three criteria were of special interest: grade point average (GPA), naval science grades (NSGs), and naval aptitude ratings (APTs). All criteria were standardized within school to a mean of 0 and a standard deviation of 1. Standardization was performed whenever at least two individuals in the sample attended a particular school. Because standardization could not be performed on any student who was the only attendee at a particular school, such cases were dropped from all analyses. For the 1982F and 1985F samples, first-year grades were cumulated. For the 1982S and 1983S samples, first- and second-year grades were cumulated.

GPA is the student's cumulative grade average from all courses. NSG is the student's cumulative grade average from all naval science courses, such as navigation and seamanship. APT is derived from NROTC staff ratings on non-academic elements of NROTC program performance, such as leadership and military bearing.

Analyses

Validity Analyses

Separate regression equations were examined by Mattson et al. (1987) to predict each of the second-year criteria for 1979-1980 entering classes (i.e., GPA, NSG, and APT). In addition, equations to predict composite criteria (GPA + APT and GPA + APT + NSG) were also examined. Unstandardized regression weights from these analyses were used to weight each of the five predictors in creating prediction equations corresponding to each of the criteria for 1982F, 1982S, and 1985F samples. For composite criteria, the regression weight for each predictor for each criterion in the composite was divided by the standard deviation of the criterion in the development sample. For example, the weight for HSR in the GPA + APT composite was calculated as follows:

$$b_{HSR*(GPA+APT)} = b_{HSR*GPA}/SD_{GPA} + b_{HSR*APT}/SD_{APT}.$$

In this equation, $b_{HSR*(GPA+APT)}$ is the unstandardized regression weight for HSR in the equation to predict the GPA+APT composite; $b_{HSR*GPA}$ is the unstandardized regression weight for HSR in the equation to predict GPA; SD_{GPA} is the standard deviation of GPA in the development sample, $b_{HSR*APT}$ is the unstandardized regression weight for HSR in the equation to predict APT; and SD_{APT} is the standard deviation of APT in the development sample. The predicted criteria were then correlated with the actual criteria to determine their cross-validity.

In the present study, regression analyses were also conducted on the 1982F, 1982S, and 1985F samples. Unstandardized regression weights from these analyses were used to create prediction equations. Weights for composite criteria were calculated as described above. Equations to predict a sixth criterion, GPA + NSG composite (combining the two academic performance criteria), were also created. Each prediction equation was cross-validated by applying it to each non-development sample. For example, the equation developed on 1982F to predict GPA was applied to both the 1982S and 1985F samples.

Obtained validity and cross-validity coefficients were corrected for predictor range restriction. Because scholarship recipients in all samples were selected on the five predictors that appear in the prediction equations under investigation, the direct restriction model was applied. Equation 10-12 in Ghiselli, Campbell, and Zedeck (1981) was used:

$$r_{xy} = \frac{r'_{xy}(\sigma_x/\sigma'_x)}{\sqrt{1 - r'^2_{xy} + r'^2_{xy}(\sigma^2_x/\sigma'^2_x)}}$$

where r_{xy} is the validity coefficient in the unrestricted sample, r'_{xy} is the validity coefficient in the restricted sample, σ_x is the standard deviation of the predictor in the unrestricted sample, and σ'_x is the standard deviation of the predictor in the restricted sample. The unrestricted sample of 1982 was all noncontract, non-BOOST, Navy-option finalists (N = 10,072). Because the contract status

was unavailable for 1985, the unrestricted 1985 sample was all non-BOOST Navy-option Finalists (N=11,968).

Standardized regression weights for each prediction equation were converted to "effective weights." These weights are constrained so that the sum of the absolute values of the weights in each equation equals 100. Effective weights provide a useful means of comparing the contribution to prediction of each predictor across samples.

ACT-SAT Conversion Tables

The ACT-SAT conversion tables were developed using test scores for those applicants from the 1985 entering class who reported both ACT and SAT scores. A total of 3,092 applicants had complete ACT and SAT scores. Separate conversion tables were developed for ACT-English to SAT-Verbal, ACT-Math to SAT-Math, and ACT-English + Math Composite to SAT-Verbal + Math Composite. Theoretically, the SAT composite scores could vary from 400 to 1600; ACT composite scores could vary from 2 to 72. For verbal and math subtests, these ranges were 200 to 800 for SAT and 1 to 36 for ACT.

Conversion table development proceeded in several stages. First, frequency distributions for the subtests and composites were constructed. A cumulative percentage for any score point (for either subtest or composite) was defined as the percentage of cases with a test score below that score point. The lowest observed ACT score was converted to the lowest observed SAT score. The percentile rank of the next lowest ACT score was then compared to the percentile ranks of the SAT scores. If the percentile rank for that ACT score was exactly equal to the percentile rank of a SAT score, the ACT score was converted directly to that SAT score. If there was no SAT score with the appropriate percentile rank, linear interpolation was used to estimate the converted SAT score. This process was repeated up through the highest observed ACT score.

Linear extrapolation was applied at the upper end of the conversion table if the highest observed ACT score was lower than the theoretical maximum (i.e., less than 36 for the ACT subtests and less than 72 for the ACT composite). Similarly, linear extrapolation was applied at the lower end of the conversion table if the lowest observed ACT score was greater than the theoretical minimum. Thus, there was a converted SAT score for each possible score on the ACT. Finally, the converted SAT scores were rounded to the traditional SAT scale so that all test scores were evenly divisible by 10.

It should be remembered that scores on the ACT subtests and composite should not be considered equivalent (or equated, in a technical sense) to corresponding scores on the SAT. Strictly speaking, test scores can be considered equivalent only if they are derived from parallel forms of the same test. This is obviously not the case for the ACT and SAT subtests and, especially, the composites. The conversion tables included here present SAT scores that are comparable to ACT scores only in the sense of having equivalent percentile ranks.

The adequacy of the conversion procedure was evaluated by computing the root mean squared difference between an applicant's observed SAT scores and the converted values as obtained from the observed ACT scores and the conversion tables. For the subtests, this evaluation was performed for the newly developed conversion table and also for the conversion table that is currently in use.

For the composites, only the newly developed conversion table was evaluated; the table currently in use does not include composite- score conversions. An additional evaluation was made for the composite-score conversions: the root mean squared difference between observed SAT scores and converted values defined as the sum of individually converted subtests was also computed. This permitted a comparison of the adequacy of converted composite scores that were defined two different ways: (1) observed ACT composite scores directly converted to the SAT scale, and (2) ACT subtests individually converted to the SAT scale and then summed.

Data from the 1985 entering class were used to develop this new conversion table. Consequently, evaluating the table using the same set of data should yield a more favorable result than is likely to occur when the table is implemented with different classes. Therefore, each conversion table was also evaluated using ACT and SAT data reported for applicants in the 1982 class.

Mean vs. Highest SAT/ACT Composite

Mean SAT or SAT-equivalent verbal and math subtests scores were computed for students in the 1983S sample.[1] Mean subtest scores were then summed to create a mean composite score for each individual. Zero-order correlations between these mean composites and all six criteria were obtained. For comparison, zero-order correlations of the highest composite were also obtained. Individuals who were tested only once were included in the analyses. Mean composite score is equal to highest composite score for these individuals (N = 371, or 48% of the sample).

New SCII and BQ Scales

The 325 items of the SCII were evaluated for possible use in selection. Each item was correlated with GPA, NSG, and APT in the 1982F and 1982S samples. Thus, there were six correlations for each item. Items were chosen for either the Grades composite (GPA + NSG), the APT criterion, or both based on the following standards: (1) at least three of the four correlations against Grades or both correlations against APT were significant ($p < .05$); and (2) endorsement rates for the different response options were relatively even. Accordingly, items in the predictor keys have reasonably consistent validities across time (first- and second-year criteria) and, for the Grades composite, across the GPA and NSG criteria. Also, they cannot easily be faked (i.e., the keyed response alternative is not obvious according to the endorsement rate findings).

Using these criteria, 56 items were selected to form a scoring key to predict Grades (GPA + NSG), and 35 items were selected to form a scoring key to predict APT. These keys are labeled SCGRADE and SCAPT, respectively. The keyed items were given positive or negative weights based on the direction of relationship exhibited with the criterior.

Similar analyses were conducted for the 34 BQ items using the 1982F and 1982S samples. The first 13 items of the BQ have five response options; the remaining 21 items have two response

---

[1]SAT and ACT data appear on different data tapes than the Highest Composite variable. In 72 cases, individuals with a Highest Composite entry did not have complete data on the SAT and/or ACT tapes. These cases were excluded from the mean-versus-highest-composite analyses.

options. For all items, Pearson correlations were computed, as well as mean criterion scores for those students endorsing each alternative. In addition, for items with more than two alternatives, eta coefficients were computed. Standards for item selection closely paralleled the standards used for the SCII scales. Using these standards, five items were selected for predicting Grades, and seven items were selected for predicting APT. These new scoring keys were labeled BQGRADE and BQAPT, respectively. Each keyed item was given a positive or negative weight based on its relationship with the criterion of interest.

Next, the incremental validities of the new grades (SCGRADE, BQGRADE) and aptitude (SCAPT, BQAPT) keys, when combined with SAT, HSR, and INTVIEW predictors, were computed on the 1983S sample. For comparison, the incremental validities of the current keys (SCRET, BQRET) were also computed.

RESULTS

Zero-order validities for individual predictors are contained in Table 3. In all samples, HSR has the highest zero-order validity with GPA, NSG, and APT. Predictor intercorrelations are reported in Appendix A. The predictor intercorrelations range from -.223 to +.194. Criterion intercorrelations are presented in Appendix B. GPA, NSG, and APT are moderately intercorrelated ($r = .41$ to .62). Composite criteria (i.e., GPA + APT, GPA + APT + NSG, and GPA + NSG) are highly intercorrelated ($r = .80$ to .95) as would be expected of composites with overlapping components.

Tables 4-9 present the validities and cross-validities of various prediction equations against NROTC program criteria. It is clear from the tables that validities for all of the prediction equations shrink only slightly from development to cross-validation samples. In some instances, cross-validities are actually higher than validities. Predictive validities are lowest for APT.

Restriction-of-range corrections had only slight impact on the validities and cross-validities (mean increase = .05). This is probably due to the fact that the population of interest, Finalists, is itself restricted in range. All Finalists must meet minimum SAT score requirements. Corrected validities and cross-validities are contained in Appendix C.

Effective weights are presented in Appendix D. While there is some variance in the effective weights obtained for each predictor across samples, the two predictors with the largest effective weights remain constant for both GPA and NSG. For both of these criteria, HSR has the largest effective weight and SAT has the second largest. For APT, HSR again has the largest effective weight, and BQRET has the second largest effective weight in three of four samples. It is clear that moderate changes in the weighting of predictors has little effect on the multiple correlations obtained.

Tables 10 and 11 present the new and current conversion tables for verbal and math subtests, respectively. Table 12 presents the new conversion table for the composite score. Table 13 presents the root mean squared errors between observed SAT scores and the estimated values obtained using the various conversion tables. The same pattern of results was obtained with the 1982 evaluation sample as was obtained with the 1985 development sample. That is, the new conversion tables for the individual subtests have errors that are slightly but consistently smaller than the errors

for the conversion tables currently in use. This difference between the new tables and the current tables remains (and is slightly larger) when the individual subtests are first converted and then summed to form a composite; that is, the new tables have smaller errors than the tables currently in use. When only the new tables are considered, it makes no difference whether the subtest scores are converted to the SAT scale and then summed or whether they are first summed and then directly converted to the SAT scales. The errors for the converted composite scores are essentially identical for these two methods.

Table 3

Zero-order Validities

Criterion	Predictor SAT	HSR	INTVIEW	SCRET	BQRET
1982F (N = 1452)					
GPA	.161**	.268**	-.001	-.060	.002
APT	.024	.129**	.054	.059	.085**
NSG	.120**	.167**	.049	.016	.056
GPA + APT	.109**	.234**	.032	-.001	.052
GPA + APT + NSG	.126**	.233**	.042	.006	.059
GPA + NSG	.159**	.247**	.028	-.026	.033
1982S (N = 1202)					
GPA	.155**	.258**	-.008	-.080*	-.025
APT	.035	.153**	.055	.022	.067*
NSG	.119**	.195**	.014	.004	.044
GPA + APT	.113**	.245**	.028	-.034	.025
GPA + APT + NSG	.126**	.247**	.025	-.022	.035
GPA + NSG	.152**	.252**	.004	-.043	.011
1985F (N = 1329)					
GPA	.144**	.269**	.021	-.067*	.003
APT	.006	.102**	.065*	.036	.103**
NSG	.142**	.195**	.007	.012	.071*
GPA + APT	.089**	.220**	.051	-.018	.063
GPA + APT + NSG	.122**	.236**	.039	-.008	.074*
GPA + NSG	.162**	.263**	.016	-.031	.042

*$p \leq .01$.
**$p \leq .001$.

Table 4

Validities and Cross-validities for GPA Predictor Composite on GPA Criterion

	Development Sample			
Cross-validation Sample	1979-80	1982F	1982S	1985F
1979-80	.315[a]	---	---	---
1982F	.284	.293	.291	.290
1982S	.278	.279	.281	.276
1985F	.277	.289	.287	.292

Note. All correlations are significant at p ≤ .001. Underlined values are the validity coefficients; all other values are cross-validity coefficients.

[a]From Mattson, Neumann, and Abrahams (1987).

Table 5

Validities and Cross-validities for APT Predictor Composite on APT Criterion

	Development Sample			
Cross-validation Sample	1979-80	1982F	1982S	1985F
1979-80	.204[a]	---	---	---
1982F	.145	.178	.172	.169
1982S	.154	.175	.180	.169
1985F	.112	.157	.154	.165

Note. All correlations are significant at p ≤ .001. Underlined values are the validity coefficients; all other values are cross-validity coefficients.

[a]From Mattson, Neumann, and Abrahams (1987).

Table 6

Validities and Cross-validities for NSG Predictor Composite on NSG Criterion

	Development Sample			
Cross-validation Sample	1979-80	1982F	1982S	1985F
1979-80	.209[a]	---	---	---
982F	.202	.215	.210	.208
1982S	.216	.218	.223	.221
1985F	.230	.245	.248	.250

Note. All correlations are significant at p ≤ .001. Underlined values are the validity coefficients; all other values are cross-validity coefficients.

[a]From Mattson, Neumann, and Abrahams (1987).

Table 7

Validities and Cross-validities for GPA + APT Predictor Composite on GPA + APT Criterion Composite

	Development Sample			
Cross-validation Sample	1979-80	1982F	1982S	1985F
1982F	.236	.257	.254	.251
1982S	.245	.254	.258	.251
1985F	.215	.242	.241	.248

Note. All correlations are significant at p ≤ .001. Underlined values are the validity coefficients; all other values are cross-validity coefficients.

Table 8

Validities and Cross-validities for GPA+APT+NSG Predictor Composite on GPA+APT+NSG Criterion Composite

	Development Sample			
Cross-validation Sample	1979-80	1982F	1982S	1985F
1982F	.246	.267	.263	.264
1982S	.254	.262	.266	.261
1985F	.243	.270	.269	.274

Note. All correlations are significant at $p \leq .001$. Underlined values are the validity coefficients; all other values are cross-validity coefficients.

Table 9

Validities and Cross-validities for GPA+NSG Predictor Composite on GPA+NSG Criterion Composite

	Development Sample			
Cross-validation Sample	1979-80	1982F	1982S	1985F
1982F	.270	.282	.279	.279
1982S	.269	.272	.274	.272
1985F	.282	.297	.297	.300

Note. All correlations are significant at $p \leq .001$. Underlined values are the validity coefficients; all other values are cross-validity coefficients.

Table 10

ACT to SAT Conversion Tables: Verbal

Original ACT Score	Converted SAT Score PDRI	Converted SAT Score Current
36	800	750
35	770	750
34	740	750
33	710	750
32	710	740
31	700	720
30	690	690
29	660	670
28	640	650
27	620	630
26	590	610
25	570	590
24	540	570
23	510	540
22	490	510
21	460	490
20	440	460
19	420	440
18	410	420
17	390	400
16	380	380
15	370	370
14	350	350
13	330	340
12	320	320
11	290	310
10	270	300
9	260	280
8	260	270
7	250	250
6	250	240
5	230	230
4	230	220
3	220	220
2	210	210
1	200	200

Table 11

ACT to SAT Conversion Tables: Math

Original ACT Score	Converted SAT Score PDRI	Converted SAT Score Current
36	740	780
35	730	750
34	710	730
33	690	710
32	680	690
31	670	680
30	650	670
29	630	640
28	610	620
27	590	590
26	570	560
25	540	530
24	520	520
23	510	500
22	490	490
21	470	480
20	450	470
19	440	460
18	430	440
17	420	430
16	410	410
15	410	390
14	400	370
13	380	360
12	360	350
11	350	340
10	340	330
9	340	320
8	320	310
7	310	300
6	290	290
5	280	280
4	270	270
3	240	270
2	200	260
1	200	200

Table 12

ACT to SAT Conversion Table: Verbal + Math Composite Scores Directly Converted

Original ACT Score	Converted SAT Score
72	1600
71	1560
70	1510
69	1470
68	1460
67	1460
66	1440
65	1420
64	1410
63	1390
62	1350
61	1340
60	1320
59	1290
58	1270
57	1260
56	1240
55	1210
54	1190
53	1170
52	1160
51	1140
50	1110
49	1090
48	1070
47	1050
46	1020
45	1000
44	980
43	960
42	950
41	930
40	910
39	890
38	870

Table 12 (Continued)

Original ACT Score	Converted SAT Score
37	860
36	840
35	820
34	810
33	800
32	790
31	780
30	770
29	760
28	750
27	740
26	730
25	720
24	700
23	680
22	670
21	620
20	600
19	600
18	590
17	580
16	570
15	570
14	570
13	540
12	490
11	440
10	440
9	430
8	430
7	420
6	420
5	420
4	410
3	410
2	400

Table 13

Root Mean Squared Error Between Observed SAT Score and the Estimated Value Obtained Using a Conversion Table

	Conversion Table						
	Individual Subtests				Composites		
	Verbal		Math		Sum of Individually Converted Subtests		
Entering Year	Current	PDRI	Current	PDRI	Current	PDRI	Directly Converted
1982 (N=3122)	72.98	72.16	59.70	57.15	102.02	99.04	98.51
1985 (N=3092)	71.14	69.36	61.07	57.71	98.47	94.93	95.48

Note. The conversion tables were developed using data from the 1985 entering class.

Use of the highest SAT/ACT composite generally results in higher validities against college performance criteria than does use of the mean composite. The improvement in validities is minimal, however. Table 14 presents these uncorrected validities. The correlation between average and highest composite is.92.

Table 14

Uncorrected Validities of Highest and Mean SAT or SAT-equivalent Composite

	Predictor	
Criterion	High Composite	Mean Composite
GPA	.158**	.142**
APT	.040	.026
NSG	.142**	.140**
GPA+APT	.113**	.095*
GPA+APT+NSG	.137**	.123**
GPA+NSG	.168**	.158**

Note. 1983S sample with complete SAT/ACT data (N=773).

*p < .01.
**p < .001.

Incremental validities of the newly developed SCII and BQ grades and aptitude scoring keys (SCGRADE, BQGRADE, SCAPT, and BQAPT) are presented in Table 15. Incremental validities of the current SCRET and BQRET predictors are included for comparison. In general, the new grades keys add more to the prediction of the criteria than do the aptitude keys or the current keys. None of the keys significantly improves prediction of the NSG criterion, however. The slight superiority of the new SCGRADE and BQGRADE keys over the current SCRET and BQRET keys is to be expected; the current keys were designed to predict career retention rather than NROTC performance.

Table 15

Incremental Validities for Old and New SCII and BQ Scoring Keys

Criterion	Variables in Equation	Multiple R	R^2	Change in R^2 from SAT, HSR, INTVIEW Equation	
GPA					
	SAT, HSR, INTVIEW	.291	.085	---	
	+BQRET, SCRET	.293	.086	.001	
	+BQGRADE, SCGRADE	.308	.095	.010*	
	+BQAPT, SCAPT	.292	.085	.001	
APT					
	SAT, HSR, INTVIEW	.214	.046	---	
	+BQRET, SCRET	.225	.051	.005	
	+BQGRADE, SCGRADE	.237	.056	.011*	
	+BQAPT, SCAPT	.233	.054	.009	
NSG					
	SAT, HSR, INTVIEW	.258	.066	---	
	+BQRET, SCRET	.277	.076	.010	(p=.011)
	+BQGRADE, SCGRADE	.267	.071	.005	
	+BQAPT, SCAPT	.266	.071	.004	
GPA+APT					
	SAT, HSR, INTVIEW	.280	.078	---	
	+BQRET, SCRET	.286	.082	.003	
	+BQGRADE, SCGRADE	.303	.092	.013*	
	+BQAPT, SCAPT	.285	.081	.003	
GPA+APT+NSG					
	SAT, HSR, INTVIEW	.299	.089	---	
	+BQRET, SCRET	.309	.095	.006	
	+BQGRADE, SCGRADE	.316	.100	.011*	
	+BQAPT, SCAPT	.304	.092	.003	
GPA+NSG					
	SAT, HSR, INTVIEW	.307	.094	---	
	+BQRET, SCRET	.315	.099	.005	
	+BQGRADE, SCGRADE	.319	.102	.007	
	+BQAPT, SCAPT	.308	.095	.001	

Note. 1983S sample (N=845).
*$p < .01$.

DISCUSSION AND CONCLUSIONS

The validity of the QI for predicting success in NROTC is stable across different student cohorts (i.e., 1979-1985) and across freshman and sophomore criterion years. The effective weights for the predictors do change somewhat across time and criterion years, however. Modifications of some QI components seemed warranted.

The newly developed ACT-SAT subtest conversion tables result in smaller root mean squared errors than do the current tables. Use of these new conversion tables may improve the validity of SAT. It does not seem to matter empirically whether composite scores are directly converted or whether the individual subtests are first converted and then summed to form a composite score.

Highest SAT or SAT-equivalent composite results in slightly greater validity against college performance criteria than does the average composite (mean difference = .012). This result is discrepant with previous research, however. In a study of Naval Academy applicants, Cowen and Abrahams (1981) found average college aptitude test score had significantly higher validity than highest college aptitude test score against first year academic performance. Boldt, Centra, and Courtney (1986) investigated the validities of four methods of treating multiple SAT scores: the highest verbal + math (these may be from separate administrations), the most recent score, the score from the one administration with the highest verbal+math, and the average score. Boldt et al. (1986) found the average to be slightly better than the other three methods of treating multiple scores, with an increase in validity of .01 to .02. This validity increase is much smaller than the increase found by Cowen and Abrahams (1981) (.103). Boldt et al. (1986) advocate using the average score when validity is the primary concern because the average is simple to use and is more valid than using the highest or latest scores. In light of previous research and the concern that use of the highest composite may disadvantage lower income and racial minority students, who tend to have a lower rate of multiple testings (Boldt et al., 1986), use of the average score is preferable.

HSR is the most heavily weighted component. Consequently, the failure to have an applicant's HSR creates a serious deficiency in evaluating candidates. Records from the 1985 sample indicate that approximately 12 percent of Finalists are missing the information necessary to compute HSR.

An alternative for dealing with such cases is to estimate HSR statistically. One procedure, developed by Alf (1987), uses either SAT or ACT scores for estimating HSR. This procedure results, effectively, in increasing the SAT/ACT weights in the QI for those Finalists whose HSR is missing. This essentially requires the use of two different sets of weights in computing the QI, one set for Finalists with HSR and one set for Finalists without HSR. While preliminary analyses with this estimation procedure reveal slight improvement in the prediction of college performance (Alf, 1987), the increase appears too small to justify application of this procedure.

A second alternative is to consider the use of an independently measured surrogate, such as self-reported high school performance. Preliminary comparisons between actual and self-reported grades have shown high correlations between the two measures as well as substantial validities for self-reported grades that are largely independent of SAT/ACT (Neumann & Abrahams, personal communication, August 1987). These results clearly suggest that self-reported HSR should be explored for those Finalists with missing officially reported HSR.

The interview (INT) predictor needs improvement. Its validities and effective weights for most criteria are small, with the exception of the APT criterion. The values of this predictor are skewed, with most Finalists receiving the highest possible rating. Structured interview procedures that probe for past behavior and accomplishments related as closely as possible to the criteria of interest (e.g., responsibility, leadership) have been more successful than unstructured interviews in predicting future performance (Janz, 1982; Latham, Saari, Pursell, & Campion, 1980). Thus a structured, construct-oriented interview approach with descriptive anchors for high, mid-range, and low predicted performance for each construct (e.g., leadership) rated should result in a broader range of criteria addressed and better validity.

Newly developed SCII and BQ keys (SCGRADE and BQGRADE) designed to predict grades provide slightly greater incremental validity when combined with SAT, HSR, and INTVIEW predictors than do the current retention keys (SCRET and BQRET). Before adoption of the new keys, their relationship to other important criteria, such as retention, should be investigated.

Because some items on the current background questionnaire are outdated and/or sexist, future research should include the development and validation of a new background questionnaire.

The focus of this investigation has been on college performance and NROTC performance criteria. Other important dependent measures, such as attrition, should be studied. Further research should investigate means of predicting different types of NROTC attrition and designing interventions to decrease the attrition rate.

RECOMMENDATIONS

1. A close approximation of the QI should be used to aid in the selection of NROTC scholarship recipients.

2. The new ACT/SAT conversion tables should replace the tables currently in use.

3. Although highest SAT/ACT score is slightly more valid than average score, previous research and other concerns (e.g., equity) warrant the use of the average score for selection of Navy ROTC applicants.

4. The value of using self-reported high school rank to replace missing data on this variable should be investigated.

5. A new structured officer interview should be developed.

6. Although the new SCII and BQ keys provide slightly greater incremental validity than the existing keys, no recommendation is warranted until further research determines their relationship to retention.

7. A new BQ should be developed and validated.

REFERENCES

Alf, E. (17 March 1987). *Estimating missing rank in class data* (NPRDC Letter Report). San Diego: Navy Personnel Research and Development Center.

Boldt, R. F., Centra, J. A., & Courtney, R. G. (1986). *The validity of various methods o treating multiple SAT scores* (College Board Report No. 86-4). New York: College Entrance Examination Board.

Cowen, M. B., & Abrahams, N. M. (1981). *Selecting qualified candidates to the United States Naval Academy using college aptitude test scores* (NPRDC Special Report 82-20). San Diego: Navy Personnel Research and Development Center.

Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences.* San Francisco: W. H. Freeman & Co.

Janz, T. (1982). Initial comparisons of patterned behavior description interviews versus unstructured interviews. *Journal of Applied Psychology, 67*, 577-580.

Latham, G. P., Saari, L. M., Pursell, E. D., & Campion, M. A. (1980). The situational interview. *Journal of Applied Psychology, 65*, 422-427.

Mattson, J. D., Neumann, I., & Abrahams, N. M. (1987). *Development of a revised composite for NROTC selection* (MPL Technical Note 87-7). San Diego: Navy Personnel Research and Development Center.

Neumann, I. (December 1978). *NROTC scholarship program conversion table: ACT scores to CEEB scores* (NPRDC Letter Report). San Diego: Navy Personnel Research and Development Center.

Neumann, I., & Abrahams, N. M. (February 1978). *Construction and validation of a Strong Campbell Interest Inventory Career Tenure Scale for use in selecting NROTC midshipmen* (NPRDC Letter Report). San Diego: Navy Personnel Research and Develcpment Center.

APPENDIX A

PREDICTOR INTERCORRELATIONS

Table A-1

Predictor Intercorrelations

Predictor	SAT	SR	INTVIEW	SCRET
1982F (N = 1452)				
SAT	---			
HSR	.166**	---		
INTVIEW	-.030	-.007	---	
SCRET	-.223**	-.113**	.025	---
BQRET	-.130	.003	.059	.092**
1982S (N = 1202)				
SAT	---			
HSR	.194**	---		
INTVIEW	-.028	.004	---	
SCRET	-.216**	-.102**	.002	---
BQRET	-.138**	-.014	.054	.110**
1985F (N = 1329)				
SAT	---			
HSR	.160**	---		
INTVIEW	-.022	.006	---	
SCRET	-.179**	-.091**	-.015	---
BQRET	-.137**	-.082**	.065*	.129*

*p ≤ .01.
**p ≤ .001.

APPENDIX B

CRITERION INTERCORRELATIONS

Table B-1

Criterion Intercorrelations

Criterion	GPA	APT	NSG	GPA + APT	GPA + APT + NSG
			1982F (N = 1452)		
GPA	---				
APT	.434	---			
NSG	.548	.446	---		
GPA + APT	.847	.847	.587	---	
GPA + APT + NSG	.819	.777	.824	.942	---
GPA + NSG	.880	.500	.880	.815	.934
			1982S (N = 1202)		
GPA	---				
APT	.410	---			
NSG	.616	.488	---		
GPA + APT	.840	.840	.657	---	
GPA + APT + NSG	.825	.773	.857	.952	---
GPA + NSG	.899	.500	.899	.833	.936
			1985F (N = 1329)		
GPA	---				
APT	.416	---			
NSG	.560	.408	---		
GPA + APT	.841	.841	.576	---	
GPA + APT + NSG	.823	.759	.820	.940	---
GPA + NSG	.883	.466	.883	.802	.930

Note. All correlations are significant at $p \leq .001$.

APPENDIX C

VALIDITIES AND CROSS-VALIDITIES CORRECTED FOR PREDICTOR RANGE RESTRICTION

Table C-1

Validities and Cross-Validities for GPA Predictor Composite on GPA Criterion, Corrected for Predictor Range Restriction

	Development Sample			
Cross-validation Sample	1979-80	1982F	1982S	1985F
1982F	.325	<u>.345</u>	.336	.342
1982S	.312	.322	<u>.319</u>	.321
1985F	.323	.345	.339	<u>.355</u>

Note. Underlined values are the validity coefficients; all other values are cross-validity coefficients.

Table C-2

Validities and Cross-Validities for APT Predictor Composite on APT Criterion, Corrected for Predictor Range Restriction

	Development Sample			
Cross-validation Sample	1979-80	1982F	1982S	1985F
1982F	.167	<u>.216</u>	.212	.209
1982S	.175	.210	<u>.220</u>	.208
1985F	.136	.201	.204	<u>.220</u>

Note. Underlined values are the validity coefficients; all other values are cross-validity coefficients.

Table C-3

Validities and Cross-Validities for NSG Predictor Composite on NSG Criterion, Corrected for Predictor Range Restriction

	Development Sample			
Cross-validation Sample	1979-80	1982F	1982S	1985F
1982F	.246	<u>.269</u>	.255	.252
1982S	.258	.266	<u>.266</u>	.262
1985F	.289	.307	.305	<u>.306</u>

Note. Underlined values are the validity coefficients; all other values are cross-validity coefficients.

Table C-4

Validities and Cross-Validities for GPA + APT Predictor Composite on GPA + APT Criterion, Corrected for Predictor Range Restriction

	Development Sample			
Cross-validation Sample	1979-80	1982F	1982S	1985F
1982F	.276	<u>.312</u>	.303	.306
1982S	.282	.304	<u>.304</u>	.303
1985F	.265	.309	.306	<u>.326</u>

Note. Underlined values are the validity coefficients; all other values are cross-validity coefficients.

Table C-5

Validities and Cross-Validities for GPA + APT + NSG Predictor Composite on GPA + APT + NSG Criterion, Corrected for Predictor Range Restriction

	Development Sample			
Cross-validation Sample	1979-80	1982F	1982S	1985F
1982F	.292	<u>.327</u>	.317	.320
1982S	.295	.315	<u>.314</u>	.313
1985F	.302	.345	.340	<u>.352</u>

Note. Underlined values are the validity coefficients; all other values are cross-validity coefficients.

Table C-6

Validities and Cross-Validities for GPA + NSG Predictor Composite on GPA + NSG Criterion, Corrected for Predictor Range Restriction

	Development Sample			
Cross-validation Sample	1979-80	1982F	1982S	1985F
1982F	.318	<u>.341</u>	.330	.334
1982S	.310	.322	<u>.319</u>	.320
1985F	.344	.370	.365	<u>.375</u>

Note. Underlined values are the validity coefficients; all other values are cross-validity coefficients.

APPENDIX D

EFFECTIVE WEIGHTS FOR CRITERIA BY PREDICTORS AND YEAR

Table D-1

Effective Weights and Multiple Correlations for GPA by Predictor and Year

Year	Predictor SAT	HSR	INTVIEW	SCRET	BQRET	Multiple R
1979-80[a]	20	55	1	-7	-16	.315
1982F	30	62	1	-2	5	.293
1982S	27	62	-1	-9	-1	.281
1985F	23	57	4	-7	9	.292

Note. All multiple correlations are significant at $p < .001$.
[a]Entries in this row are from Mattson, Neumann, and Abrahams (1987).

Table D-2

Effective Weights and Multiple Correlations for APT by Predictor and Year

Year	Predictor SAT	HSR	INTVIEW	SCRET	BQRET	Multiple R
1979-80[a]	-11	54	12	18	-5	.204
1982F	8	36	14	20	22	.178
1982S	7	4[illegible]	16	11	20	.180
1985F	3	35	18	11	33	.165

Note. All multiple correlations are significant at $p < .001$.
[a]Entries in this row are from Mattson, Neumann, and Abrahams (1987).

Table D-3

Effective Weights and Multiple Correlations for NSG by Predictor and Year

Year	Predictor SAT	HSR	INTVIEW	SCRET	BQRET	Multiple R
1979-80[a]	19	56	10	9	5	.209
1982F	27	36	11	12	14	.215
1982S	26	47	3	10	14	.223
1985F	29	40	1	9	21	.250

Note. All multiple correlations are significant at $p < .001$.
[a]Entries in this row are from Mattson, Neumann, and Abrahams (1987).

Table D-4

Effective Weights and Multiple Correlations for GPA + APT by Predictor and Year

Year	SAT	HSR	INTVIEW	SCRET	BQRET	Multiple R
	Predictor					
1979-80[a]	8	66	8	5	-14	---
1982F	20	51	7	9	13	.257
1982S	20	62	7	1	10	.258
1985F	16	51	11	1	21	.248

Note. All multiple correlations are significant at $p < .001$.
[a]Entries in this row are from Mattson, Neumann, and Abrahams (1987).

Table D-5

Effective Weights and Multiple Correlations for GPA + APT + NSG by Predictor and Year

Year	SAT	HSR	INTVIEW	SCRET	BQRET	Multiple R
	Predictor					
1979-80[a]	12	65	9	6	-8	---
1982F	23	45	9	10	13	.267
1982S	22	56	6	4	12	.266
1985F	21	47	7	4	21	.274

Note. All multiple correlations are significant at $p < .001$.
[a]Entries in this row are from Mattson, Neumann, and Abrahams (1987).

Table D-6

Effective Weights and Multiple Correlations for GPA + NSG by Predictor and Year

Year	SAT	HSR	INTVIEW	SCRET	BQRET	Multiple R
	Predictor					
1982F	29	49	6	6	10	.282
1982S	29	61	1	1	8	.275
1985F	28	52	3	1	16	.300

Note. All multiple correlations are significant at $p < .001$.

DISTRIBUTION LIST

Distribution:
Office of Naval Technology (Code 222)
Commanding Officer, Naval Education and Training Program Management Support Activity (Code N-31) (2)
Defense Technical Information Center (DTIC) (2)

Copy to:
Chief of Naval Operations (OP-01B2)
Director, Total Force Training and Education Policy Division (OP-114) (2)
Commander, U.S. ARI, Behavioral and Social Sciences, Alexandria, VA (PERI-POT-I) (2)
Assistant for Manpower Personnel and Training Research and Development (OP-01B2)